New Ranking Methods with Application to Sports.

Anjela Govan and Carl Meyer

Department of Mathematics, Box 8205, NC State University, Raleigh, NC 27695-8205
{aygovan,meyer}@math.ncsu.edu

Our goals ...

- Generalize ranking methods PageRank and HITS to create new methods.
- Test generalized methods on National Football League data.

New Methods ...

- New method GeM (Generalized Markov) is a generalized version of Google's ranking algorithm PageRank.

Google

- New method mHITS is a modified version of a ranking algorithm HITS used in Ask search engine (mHITS developed with Amy Langville and Luke Ingram).

Ask.com

Example ...

- Selected games form 2005 NFL regular season.

Game	Team	Score	Total Yards	Team	Score	Total Yards
1	Car	27	326	NO	10	277
2	Car	34	287	TB	14	270
3	Chi	13	258	Car	3	238
4	Chi	13	239	TB	10	275
5	NO	23	291	Car	20	350
6	TB	20	247	Car	10	247
7	TB	27	285	NO	13	306
8	Pit	21	363	Chi	9	268

- Use both score differences and yard differences for GeM example.
- Use scores for mHITS example.

GeM Method

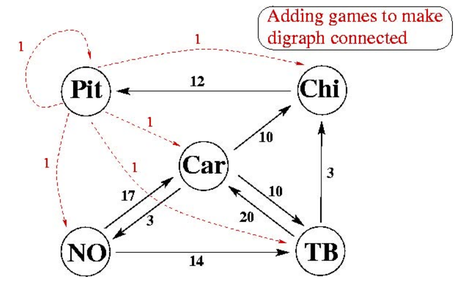
$$\mathbf{G} = \alpha_0\mathbf{S}_0 + \alpha_1\mathbf{S}_1 + \ldots + \alpha_p\mathbf{S}_p$$

- Each $\mathbf{S}_i$ is stochastic, $0 \leq \alpha_i \leq 1$ and $\Sigma\alpha = 1$.
- The vector $\boldsymbol{\pi}$ contains the rating scores of each team, such that

$$\boldsymbol{\pi}^T = \boldsymbol{\pi}^T\mathbf{G}$$

Example: one way to form $\mathbf{S}_0, \ldots, \mathbf{S}_p$ use game scores for matrix $\mathbf{S}_0$.

- Team → node in a directed graph.
- Each game → edge from loser to winner.
- Weight on each edge is the statistic difference (e.g. total yards, etc).



	Car	Chi	NO	Pit	TB
Car	0	10/23	3/23	0	10/23
Chi	0	0	0	1	0
NO	17/31	0	0	0	14/31
Pit	1/5	1/5	1/5	1/5	1/5
TB	20/23	3/23	0	0	0

- Divide score differences by each team's total score difference.

- Use total yards to form $\mathbf{S}_1$
- $\mathbf{G} = 0.65\mathbf{S}_0 + 0.35\mathbf{S}_1$
- Rating vector:
  $\boldsymbol{\pi}^T \approx \begin{pmatrix} 0.5428 & 0.4838 & 0.1955 & 0.5465 & 0.3666 \end{pmatrix}^T$
- Ranked teams: from first to last

Pit Car Cin TB NO

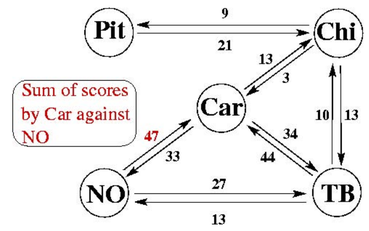
mHITS Method

$$\mathbf{d}^{(k)} = \mathbf{P}\frac{1}{\mathbf{o}^{(k)}}$$
$$\mathbf{o}^{(k)} = \mathbf{P}^T\frac{1}{\mathbf{d}^{(k-1)}}$$

- Entry p_{ij} is a statistic of team j against i.
- Vector $\mathbf{o}$ contains the offensive ratings.
- Vector $\mathbf{d}$ contains the defensive ratings.

Example: Use game scores

- Team → node in a directed graph.
- Each game → two edges between teams i, j.
- Weight on each edge from i to j is the score (or total yards, etc.) of team j against team i.



	Car	Chi	NO	Pit	TB
Car	0	13	33	0	34
Chi	3	0	0	21	10
NO	47	0	0	0	27
Pit	0	9	0	0	0
TB	44	13	13	0	0

- $\mathbf{o} \approx \begin{pmatrix} 0.40 & 177.39 & 0.25 & 0.0000000029 & 0.19 \end{pmatrix}^T$
- $\mathbf{d} \approx 1.0e+005 * \begin{pmatrix} 0.0041 & 3.012 & 0.0015 & 0.000001 & 0.0019 \end{pmatrix}^T$
- Overall ratings (normalized):
  $\mathbf{r} = \mathbf{o}(1/\mathbf{d}) \approx \begin{pmatrix} 0.1262 & 0.2044 & 0.0717 & 0.4772 & 0.1205 \end{pmatrix}^T$
- Ranked teams: from first to last

Pit Chi Car TB NO

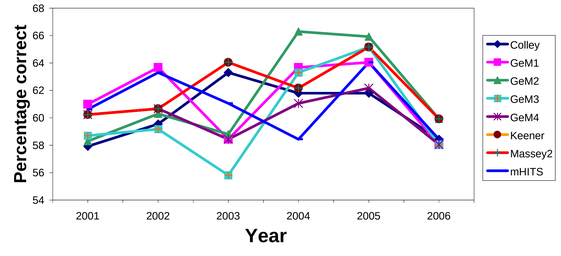
Compete against ...

- **Colley** - matrix method derived using Laplace's rule of succession (BCS computer ranking).
- **Massey** - matrix method derived using mathematical expectation (BCS computer ranking).
- **Keener** - matrix method based on Perron-Frobenius Theorem for non negative matrices.

Results ...

- **Gem1** - Basic model $\mathbf{G} = \alpha\mathbf{S}_0 + (1-\alpha)\mathbf{e}\mathbf{e}^T$
- **Gem2** - Feature vectors model
  $\mathbf{G} = \alpha_0[\mathbf{H} + \mathbf{a}\mathbf{u}^T] + \alpha_1\mathbf{e}\mathbf{v}_1^T + \ldots + \alpha_p\mathbf{e}\mathbf{v}_p^T$
  $\mathbf{v}_i$ is a probability distribution vector.
- **Gem3** - Offense-defense model
  $\mathbf{G} = \alpha[\mathbf{H} + \mathbf{a}\mathbf{u}^T] + \beta_1\mathbf{e}\mathbf{o}^T + \beta_2\mathbf{e}\mathbf{d}^T$
- **Gem4** - Feature matrices model
  $\mathbf{G} = \alpha_0[\mathbf{H} + \mathbf{a}\mathbf{u}^T] + \alpha_1\mathbf{F}_1 + \ldots + \alpha_p\mathbf{F}_p$

Results of NFL game predictions.



	2001	2002	2003	2004	2005	2006
Colley	57.92	59.55	63.3	61.8	61.8	58.43
GeM1	61	63.67	58.43	63.67	64.04	58.05
GeM2	58.3	60.3	58.8	66.29	65.92	59.93
GeM3	58.69	59.18	55.81	63.3	65.17	58.05
GeM4	60.23	60.67	58.43	61.05	62.17	58.05
Keener	60.23	60.67	64.04	62.17	65.17	59.93
Massey2	60.23	60.67	64.04	62.17	65.17	59.93
mHITS	60.62	63.3	61.05	58.43	64.04	58.43